
CHAPTER 3

The Relationship Between Neuronal Codes and Cortical Organization

BARRY J. RICHMOND

Laboratory of Neuropsychology
National Institute of Mental Health
Bethesda, Maryland

TIMOTHY J. GAWNE

Department of Physiological Optics
School of Optometry
University of Alabama
Birmingham, Alabama

The brain's function is to process sensory information, compare it to previous experience, and produce behavior. However, the brain is too complex to permit us to understand how this occurs by specifying all the elements and connections. Therefore we must try to learn what principles single neurons use to code information and how that information interacts in populations.

There are two important ideas connecting stimulus features to neuronal function and cortical organization. The first is that neurons are tuned (i.e., particular stimulus features yield big responses) and when the stimulus features are changed, the neuronal responses become smaller. The second is that neurons responding to similar stimuli or conditions are grouped together.

In the early visual system it is possible to relate one or more stimulus parameters such as edge orientation or stimulus color directly to the response strength (Hubel and Wiesel, 1962), with the stimulus eliciting the most intense response referred to as the optimal stimulus. If each neuron has an optimal feature and we take the extreme view that only the strongest response has an influence, the coding can be regarded as labeled lines, where each neuron codes for its particular optimal feature, and the pattern of activated neurons is the code. We know, however, that the responses are graded, and thus each neuron can be regarded as a filter against which the stimulus is measured (or in the motor system, the movement). When the relation between neurons is known, the message can be regarded as a vector sum of the activities of the neurons weighted by the relation among them. This approach has been used elegantly in the motor system by Georgopoulos et al. (1988) to show that motor cortex carries enough information to code for direction of movement and dynamics.

The second important property relates stimulus characteristics, neuronal responses, and the anatomical relation of neurons. Mountcastle (1957) first proposed that neurons with similar properties were grouped together in *cortical columns*. Hubel and Wiesel (1962) later showed that neurons within primary visual cortex change the direction of the peak in orientation gradually as different neurons are encountered at successive locations across the cortex.

These ideas are very powerful and have helped us to understand how sensory information is processed. However, we still do not fully understand how the activities of neurons are combined to provide the capabilities we see in behavior. Thus, we must figure out what pieces are missing.

NEURONAL CODE DIMENSIONALITY

A basic assumption of the neuron doctrine is that single neurons carry information in their responses (Barlow, 1972). For fast signaling, information is clearly carried by means of action potentials, and it is obvious that the number of action potentials is an important part of the signal. Until about 10 years ago, this was the only response parameter considered for the neural code in visual system neurophysiology. However, it is clear that the code can accommodate considerably more complexity: the arrival time of every spike, for example, can carry information.

The spike count or the average firing frequency is most often used as the response measure related to the information. When only the spike count carries information, a single number specifies the response. However, even in this simple case the response must be defined in terms of time, usually by counting the number of spikes within a specific time window.

To investigate the role of time variation, we must clarify what we are looking for. Trivially we know that many neurons have responses that are related to changes in the stimuli; for example, simple cells in V1 cycle are in phase with the rate at which a sinusoidal grating drifts across a receptive field (Movshon et al., 1977a). When stimulated by the same drifting gratings, the responses of complex cells are more constant (Movshon et al., 1977b), with the degree of constancy being dependent on spatial frequency (Spitzer and Hochstein, 1985). In general, this type of coding is a univariate function of the stimulus, here encoding some aspect of the contrast changing in time. This type of temporal modulation was termed temporal coding by Miller and Theunissen (1995).

A more complex type of coding takes place when a static stimulus is switched on and the pattern of activity changes as a function of the stimulus pattern. In this case different aspects of the stimulus can be encoded by different parameters of the response. When this type of time variation occurs, it is necessary to somehow record the spike arrival times, and thus the number of parameters needed to specify a response is increased. We could require that each spike time be explicitly represented, greatly increasing the number of parameters needed. However, we would always like to use the smallest number of parameters possible to describe the information in the response. When the pattern of the response changes in some significant manner, and the change is independent of the change in number of spikes, the response is multivariate. Miller and Theunissen called this second type of coding temporal encoding.

To account for the stimulus-related aspects of neuronal responses, it is important to estimate the dimensionality of the stimulus-related response. It is clear that temporal patterns of neural activity in the visual system change in relation to the stimulus (Fig. 3.1). However, it does not appear that the spikes in different responses occur with a precise pattern, suggesting that the response dimensionality is not as high as the number of spikes in a vigorous response.

We would like the number of parameters to be as small as possible, while representing the response components accurately enough to prevent the loss of information. Since the spikes do not appear to be exactly placed in time, there is

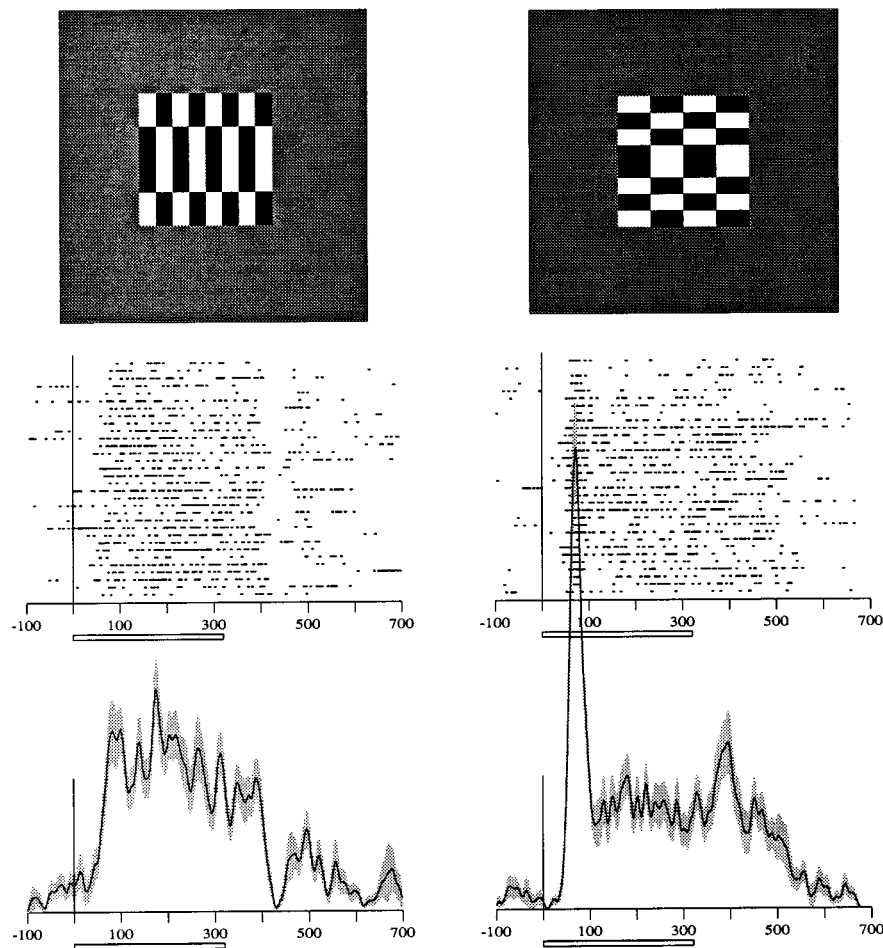


Figure 3.1 Responses of a V1 complex cell showing different patterns over time to two different stimuli. Time is represented as going from left to right. Each line of dots shows the times of the neuronal action potentials for one stimulus presentation. The stimulus appears at time zero and remains on for 320 ms, shown by the length of the hollow bar under each display. These two sets of responses have statistically indistinguishable numbers of action potentials during the stimulus presentations. However, the patterns in which the action potentials occur differ significantly.

the possibility of reducing the data to some intermediate representation, one that is richer than a single number, but still more compact than the result of specifying each spike exactly. Principal component analysis (Figs. 3.2 and 3.3) offers one systematic way to reduce the dimensionality of data (Ahmed and Rao, 1975; Richmond and Optican, 1987).

The principal components are the most efficient linear transform for a data set. Furthermore, they are ordered, with each accounting for more variance than any that follow. Thus, the first one accounts for the largest proportion of the variance. For our data, the first principal component is closely related to the spike count (Richmond and Optican, 1987), which is clearly related to the stimulus. The first principal component typically accounts for a little more (about 10% bigger) variance than the spike count, which shows that having a little temporal variation in the weighting of the counting window is helpful. When static stimuli are suddenly switched on, we find that other principal components are related to the stimulus, also (Fig. 3.3) (Richmond and Optican, 1987). In our analyses to date, up to five principal components have been needed to describe the stimulus-related aspects of the neuronal responses. This can be shown with many statistical methods including ANOVA (Richmond and Optican, 1987).

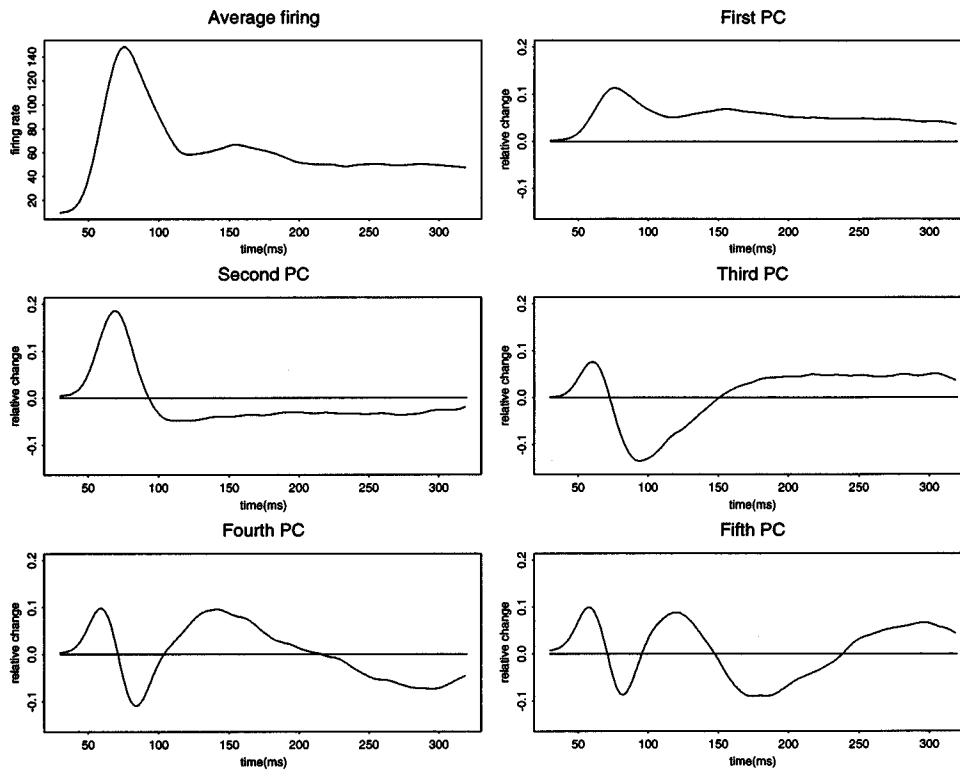


Figure 3.2 Average response and first five principal components (PC) extracted from the responses of the neuron used for Figure 3.1. The responses elicited by the whole stimulus set are used. The average response and the first principal component are very similar in shape. The subsequent PCs increase in the number of zero crossings.

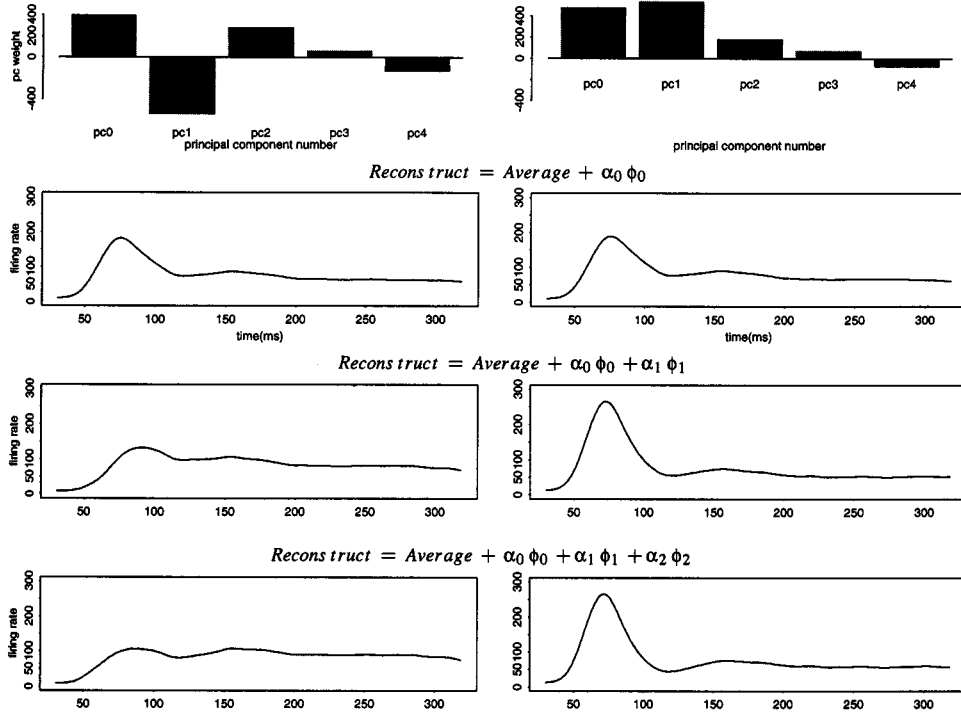


Figure 3.3 Weights of the first five principal components for 2 out of 176 stimulus patterns, 128 Walsh patterns, and 24 white and 24 black bars, from the same data shown in Figure 3.1. The weights of the first principal component are not significantly different, whereas the weights from the second are very different ($p < 0.01$, t test). The three rows of response waveforms are the response reconstructions using one, two, and three principal components, respectively.

We also want to know whether different response components carry the same or new information about the stimulus. Information theory provides a natural framework for analyzing and discussing neural signals (Optican and Richmond, 1987; Richmond and Optican, 1990). It quantifies an intuitive concept: How much information about a stimulus is carried in a spike train?

In the application of information theory, the stimuli are considered as one code and the response as another (Cover and Thomas, 1991). To apply information theory, we must postulate what the codes are, and we code the stimulus according to the question we are asking. For example, we might give each stimulus a code if we are asking how well members of a set can be distinguished. Or we could recode the stimuli according to their contrast, to allow examination of the question of whether there is more information about the stimulus identity or about the contrast. We must also postulate the response code, or we can postulate several response codes and ask which ones carry the most information.

The mutual or transmitted information is defined as

$$I(S;R) = \sum_s H(s) - H(s|r) \quad (3.1)$$

where s is the stimulus, S is the stimulus set, r is the response, R is the response set, H is the entropy or uncertainty of the stimulus, $p(s)\log p(s)$, and $H(s|r)$ is the conditional entropy about the stimulus given the response. In this form, the transmitted information is the difference between the uncertainty in the stimulus set minus the uncertainty left after a response has been received. Equation 3.1 can be rewritten as

$$I(S;R) = \sum_s \left\langle p(s|r) \log_2 \frac{p(s|r)}{p(s)} \right\rangle_r \quad (3.2)$$

where $p(s)$ is the probability of stimulus s occurring, $p(s|r)$ is the conditional probability that stimulus s elicited response r given that r was observed, and the angle brackets indicate an average over all responses r . As alluded to earlier, the response can be a vector quantity; that is, the response can be multidimensional.

As can be seen from equation 3.2, the conditional probability $p(s|r)$ must be estimated. As the response dimensionality of the response increases, estimating the probability distributions needed to calculate information becomes increasingly difficult. Consider the following example: to have no information, every response must predict each stimulus with equal probability. However, if there are so few responses that each response *could* occur once for each stimulus, then it is not possible to realize that the stimulus does not favor some responses over others. This problem has been addressed explicitly in several studies over the past several years (Optican et al., 1991; Tové et al., 1993; Kjaer et al., 1994; Panzeri and Treves, 1996; Golomb et al., 1997). At least two reasonably satisfactory procedures have been developed to make estimates of the needed probability distributions (Kjaer et al., 1994; Panzeri and Treves, 1996; Golomb et al., 1997).

In the method described by Kjaer et al., (1994), a neural network is used to estimate the conditional probabilities (Fig. 3.4). Cross-validation prevents the neural network from overlearning, hence overestimating, the information. The network gives conservative estimates of the information (i.e., the estimates tend to be low when there are not enough data). Although the computation time for the network can be long even for a univariate response, it scales well as the response dimensionality climbs. In the method described by Panzeri and Treves (1996), the amount by which the information is overestimated by the raw information measurement is itself estimated by examining the variance structure of the data. The increment of the overestimate is subtracted from the raw estimate. This method also performs well, and the computation time is short, especially for responses with few dimensions. It becomes more difficult to implement for responses with a large number of dimensions.

PRECISION OF SINGLE NEURONAL CODES

Over the past few years it has become increasingly certain that the time course of responses carries information. There are several demonstrations of time-dependent processes beyond the ones we have studied. The discoveries of Eckhorn et al. (1988) and Gray and Singer (1989) of the oscillatory periods of neural activity in the

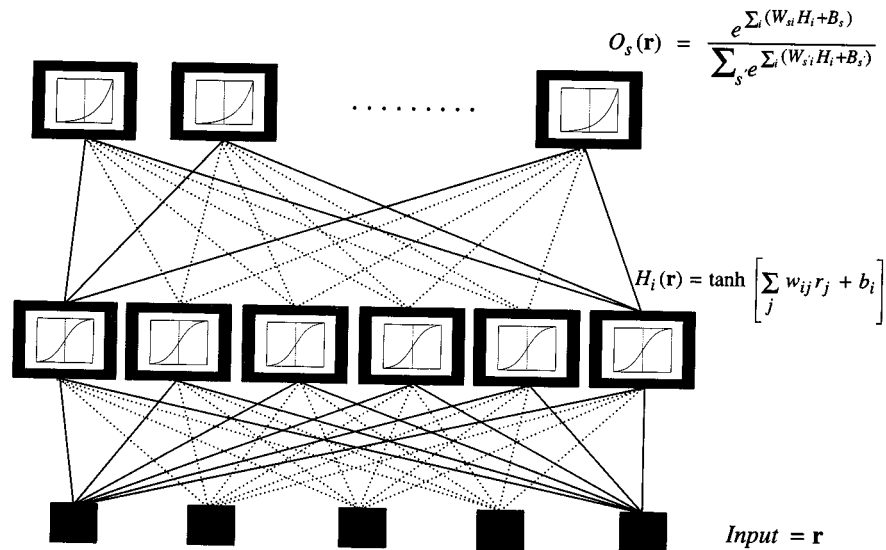


Figure 3.4 Neural network used for estimating conditional probabilities. The inputs are the response representations (e.g., spike count, principal component coefficients, spike times). The network has one hidden layer, using units having sigmoidal activation functions. The output layer has one unit for each stimulus. The activation functions are exponentials that are normalized to sum to 1.0 after each stimulus during training. During testing, the error is the negative log-likelihood, which in this classification paradigm is related only to the unit representing the current stimulus. For a complete description see Kjaer et al. (1994).

monkey visual system led to a lot of work on the role of these oscillatory periods in perception. Abeles et al. (1993) and Lestienne and Stehler (1987) have shown that there are patterns of spikes embedded in spike trains that occur substantially more frequently than chance would predict. Vaadia et al. (1995) showed that the number of repeating patterns increases in some parts of a behavioral task, suggesting that the presence of the exact patterns could be a code. All these results taken together lead us to consider what time resolution is needed to capture the stimulus-related information in the spike train. The temporal patterns we have seen change very slowly, over tens of milliseconds. The oscillations just referred to are in the neighborhood of 40 Hz, meaning that the events underlying them need be measured only with an accuracy of 25 ms or so. However, the results from Abeles and his group, and from Lestienne, suggest that there are timings accurate to the millisecond or better. And recently there have been discussions about whether some processes related to plasticity, and learning and memory, depend on the simultaneous arrival of spikes for modification of synapses (Meister, 1996; Shatz, 1996). Again, how this would work depends on time resolution—What is considered “simultaneous”?

In the analyses described in the preceding sections, the responses were low-pass-filtered by convolving the responses with a Gaussian pulse before embarking on the extraction of the principal components (Richmond and Optican, 1987). This sequence seemed to capture the information. Early in the course of our investigations, we measured the largest amount of information when the responses were

low-pass-filtered. In general, the filtering was done with cutoff frequencies in the 15–30 Hz range. In those early investigations we were not particularly concerned with defining the time resolution more carefully. Recently, we tried to detect stimulus-related information carried in response representations that used more precise time resolution. Again we made use of information theory.

In measuring information, we must postulate a code to represent the data. If more than one code could exist, we can ask which coding of the data conveys the largest amount of information. In Heller et al. (1995), we postulated 23 different codes and asked which of these carried the largest amount of stimulus-related information in data from both complex cells in V1 and single inferior temporal (IT) neurons. These codes were permutations of the exact spike times, here including the time of the first spike in the stimulus-related burst (we were fortunate in that both classes of neuron are nearly silent when no stimulus is present on the receptive field), the times of all the spikes, and a code made up of ones and zeros at each time bin (here 1 ms) of the response. We also used the spike count, the principal components of either the times, the binary code, or the low-pass-filtered response. The amounts of information carried by these different codes are shown in Figure 3.5. The best representations were principal components of low-pass-filtered responses. It always helped to include the spike count, which is nearly always a reasonably good code. In our measurements, using the neural network including the spike count as a part of the code probably gives rise to a little more information because the first principal component combines both the spike count and a little temporal weighting. If the spike count is included, the network does not need to find the spike count in the solution.

Finding that the best representation is a low-pass-filtered version of the response suggests another way to look at these results. We can ask: If we count only spikes, what is the widest spike counting window (in time) that can be used before we need to know something about the pattern of spikes to represent all the information? In Figure 3.6 we see that the information increases slowly with the width of the spike count window as we increase the window size from 8 ms to 32 ms for a V1 neuron. We found that the principal components start to have an advantage when the window is wider than 32 ms in V1 and 64 ms in IT. These effects are nearly identical to the result we got with low-pass filtering: the optimal bandwidth is about 30 Hz in V1 cortex and 15 Hz in IT cortex. These results show that very little, if any, information is missed when the response is represented by the spike count in time windows up to 30 ms.

These spike counting windows seem broad when one looks at the responses. For example, in the fly visual system the precision of the spike timing is more exact in higher frequency portions of the response (de Ruyter van Steveninck and Bialek, 1988; Bialek et al., 1991). Intuitively, this makes sense: when there are only a few spikes, the position of a spike in time can be specified broadly without affecting other spikes; however, when there are a large number of spikes, there are fewer open positions in time for them.

In our earlier work (Heller et al. 1995), we were examining all the responses taken together. Frequently there are stimuli that elicit relatively few spikes, and those influence the results because the solution for the network must cover a large range. In more recent investigations we preferentially examined the most vigorous responses. It appears that the time of the first spike in vigorous responses is more

precise; it can carry information using a time resolution of 15–20 ms. None of the other spikes in the response need to be specified at that resolution to preserve stimulus-related information. None of this precludes the use of more exact timing to carry information. Rather, we have an indication that the timing does not seem to specify the information about which stimulus appeared in this paradigm. It could also be that more exact timing is important only when considered in conjunction with other neurons.

NATURAL RESPONSE VARIABLES

The use of the principal components in conjunction with information theory led to the conclusion that the stimulus-related response was multidimensional: that is, it would take two or more parameters to describe the stimulus-related variance. The data analyses in the time domain using the time windows confirmed this result. This is a very powerful conclusion. However, the generality of the methods leaves us short of knowing much about the internal structure of the neural response code. We would like to know whether the code, even if multivariate, provides information about a single stimulus variable or, instead, whether the different response parameters encode different aspects of the stimulus.

In the past, we showed that there is independent information about luminance and stimulus form in the responses of V1 neurons (Gawne et al., 1991). At that time, we were not able to relate any specific parameter of the temporal code to a particular stimulus feature. We undertook new studies that were targeted at identifying the relation between response features and stimulus features. In these experiments we asked how single primary visual cortical V1 complex cells respond to stimulus patterns made visible by their textures rather than by their brightness. We found that the latency of complex cells doubled from 45 to 90 ms when the stimuli were made visible by texture only (Gawne et al., 1996a) (Fig. 3.7).

When we studied the relation between luminance contrast and orientation on the responses of these neurons, the latency increased as the contrast decreased (Fig. 3.8). Surprisingly, however, the latency was almost unchanged as the orientation changed at high contrast, even when the response was very weak. At low contrast both the latency and the response strength were affected by the orientation. Upon comparing the sets of data from the texture and luminance contrast experiments, we found that the patterns defined by texture had responses that were similar in latency and strength to those defined by low luminance contrast; for the textures we chose, the responses to the texture-defined stimuli were the same as the responses for bars defined by 10% luminance contrast (Fig. 3.9). Our analysis shows that the latency and the response strength are independent with respect to orientation and cue. The response strength is a code about the bar's orientation and the latency is code about how visible it is (i.e., how well the bar stands out from the background).

The classical models of complex cells cannot predict this independence of latency and response strength, since they were formulated to describe the response strength only and did not attempt to provide a description of how responses vary in time when stationary visual images appear. Therefore, we developed a model that provides these relations:

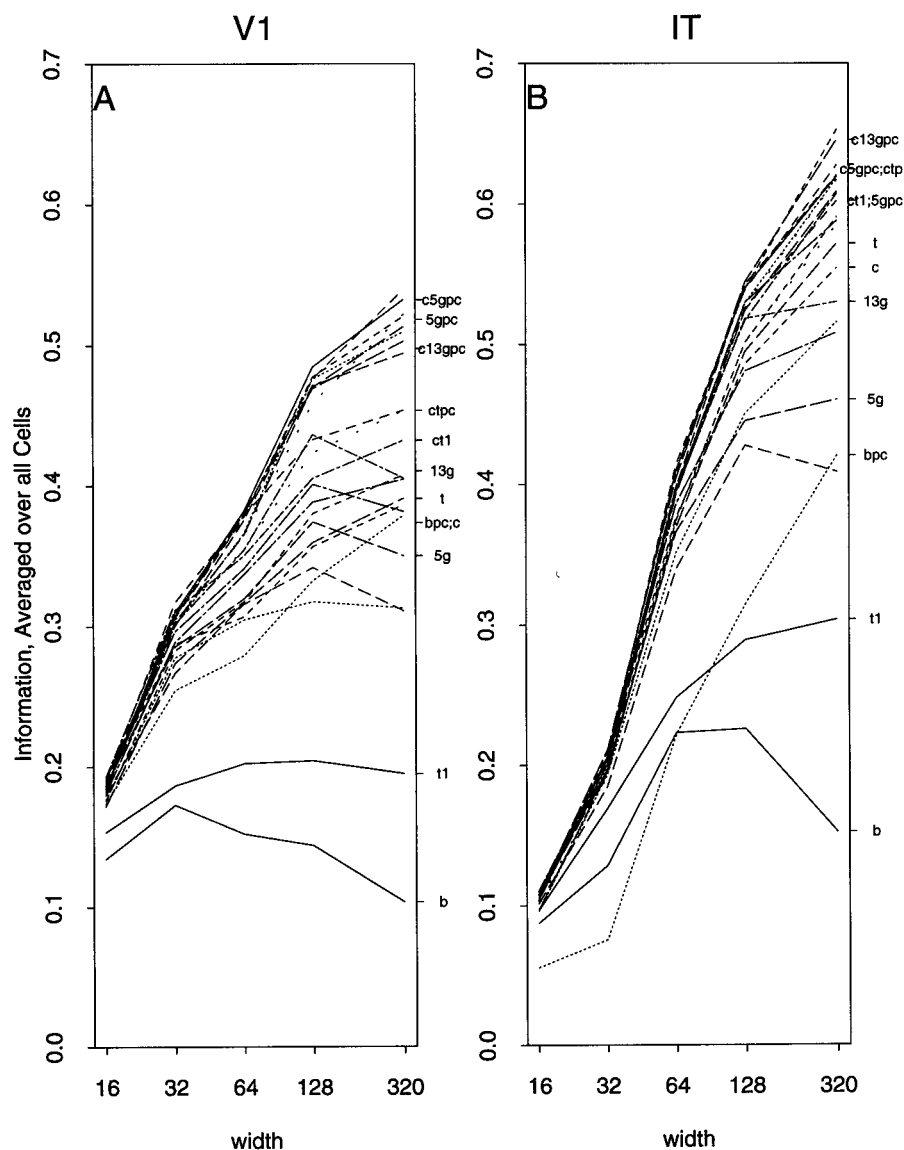


Figure 3.5 The amount of stimulus-related information that could be extracted from the responses of (A) V1 and (B) IT neurons using 23 different codes to represent the neuronal response. The representation that yielded the most information used both low-pass filtering of the spike train with a Gaussian kernel (5 ms SD in V1 and 13 ms SD in IT), followed by compression with the principal components. Generally, between three and five principal components yielded the greatest amount of information. The order of the representations for the entire 320 ms response (from best to worst) is: c5gpc c3gpc 3gpc 5gpc c8gpc 8gpc c13gpc 13gpc ctpc ct1 ct3 13g tpc ct 8g t c bpc 5g 3g t3 t1 b where c = the mean spike count; b = a binary vector in which each element represents the presence or absence of a spike in a particular 1 ms segment of the response; t = the times of the spikes; t1 = the time of the first spike; t3 = the times of the first three spikes; i = the

Instantaneous Information using PC's of Gaussians

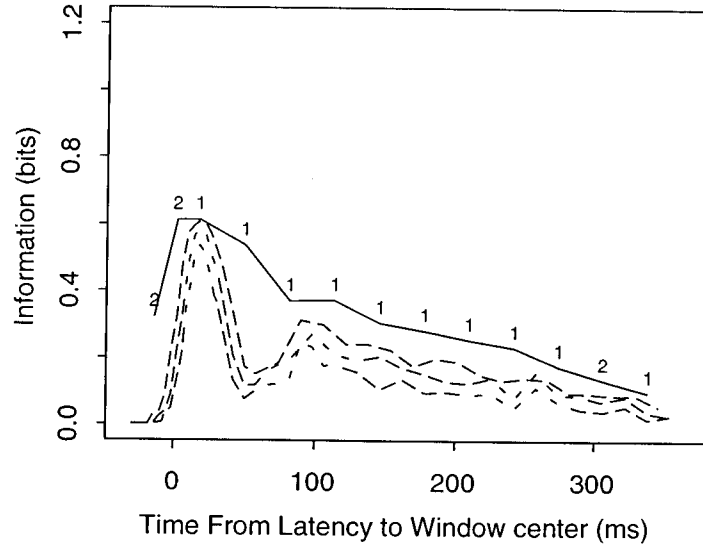
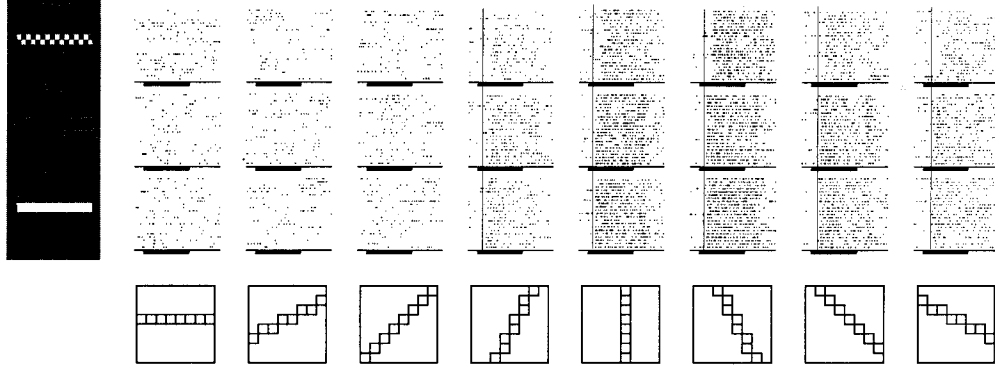


Figure 3.6 Amount of information in sliding window of different widths for one V1 neuron. Windows of different widths were slid across the responses. The stimulus-elicited information available from the spike count alone was compared to the information available using principal components. In this example the number of principal components needed to capture the information for the widest window is shown above the solid line. From the bottom, the other lines are the information in the 16, 24, 32 ms windows. In these shorter windows only one principal component was needed, and it conveyed the same amount of information as the spike count. Thus, it is only in the largest time window shown here (48 ms) that having a representation other than the spike count confers an advantage.

intervals between successive spikes; \mathbf{i}^{-1} = the reciprocals of these intervals; $\exists g \ 5g \ 8g \ 13g$ = low-pass-filtered versions of the binary impulse representation \mathbf{b} obtained by convolving the spike train with Gaussians of standard deviations of 3, 5, 8, and 13 ms, and resampling the resulting continuous waveforms at 4 ms intervals; $\mathbf{c} \ \mathbf{t}$ = the spike times as in \mathbf{t} , augmented by the spike count \mathbf{c} ; $\mathbf{c} \ \mathbf{t} \ 1$ = the first spike time and the total spike count; $\mathbf{c} \ \mathbf{t} \ 3$ = the first three spike times and the total spike count; $\mathbf{t} \ \mathbf{p} \ \mathbf{c}$ = up to five principal components of the spike times; $\mathbf{c} \ \mathbf{t} \ \mathbf{p} \ \mathbf{c}$ = principal components of the spike times, plus the spike count; $\mathbf{b} \ \mathbf{p} \ \mathbf{c} \ \exists g \ \mathbf{p} \ \mathbf{c}$ etc. = up to five principal components of the representations \mathbf{b} and $\exists g \ 5g \ 8g \ 13g$; and $\mathbf{c} \ \mathbf{b} \ \mathbf{p} \ \mathbf{c} \ \mathbf{c} \ \exists g \ \mathbf{p} \ \mathbf{c}$ etc. = $\mathbf{b} \ \mathbf{p} \ \mathbf{c} \ \exists g \ \mathbf{p} \ \mathbf{c}$ etc., augmented by the spike count.

A



B

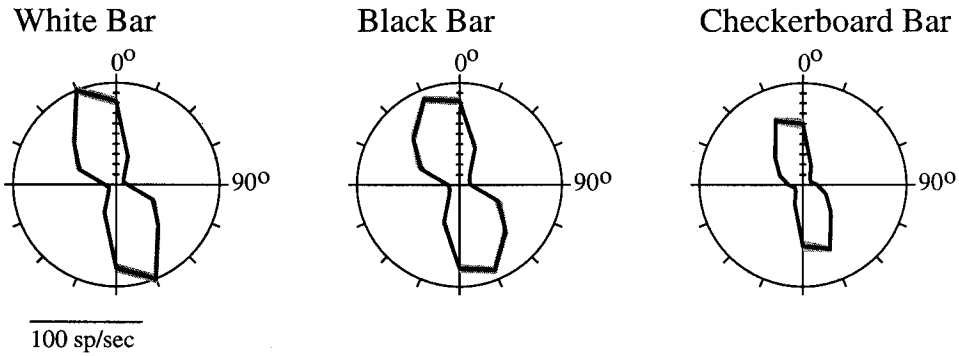


Figure 3.7 Example of the latency difference seen for a set of bars defined by texture compared to bars defined by luminance (A). The tuning curves (B) show response the number of spikes during the 270 ms stimulus presentation as the distance from the center; the direction of the point indicates the bar orientation. Tuning curves at the bottom show that the tuning remains the same. The rasters show that the response to the texture defined bar is delayed by up to 40 ms. This set was typical for the 65 complex cells recorded in these experiments.

$$\text{Strength} = a_1 \cos(\theta) + a_2 \log_{10} \alpha + a_3 \quad (3.3)$$

and

$$\text{Latency} = b_1 (2 - \log_{10} \alpha) \cos \left[2 \left(\frac{\theta - \pi}{2} \right) \right] - [b_2 \log_{10} \alpha] + b_3 \quad (3.4)$$

where a 's and b 's are constants, θ is the orientation varying between $-\pi/2$ and $\pi/2$, and α is the luminance contrast, varying between 1% and 100%. In these

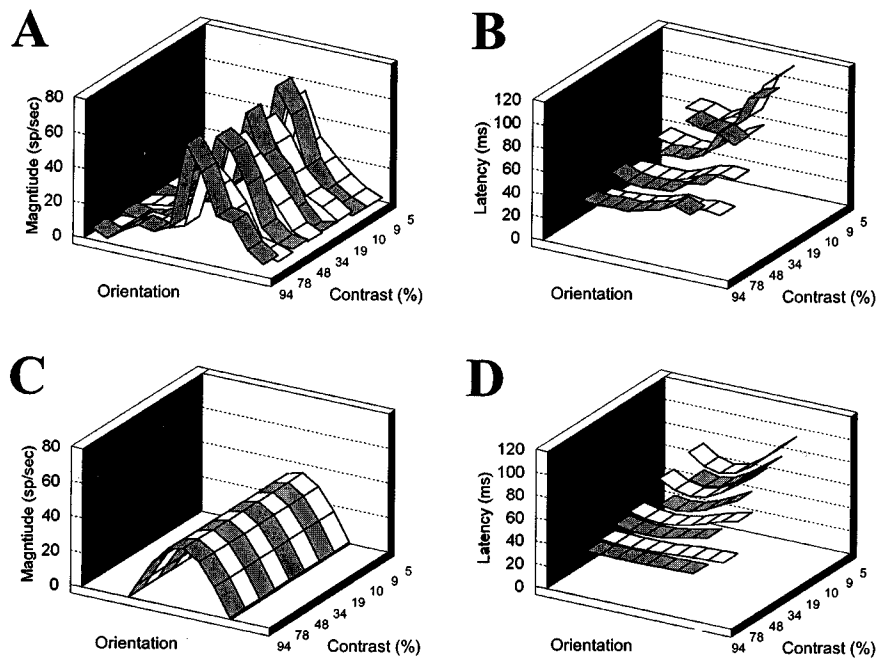


Figure 3.8 Surfaces showing the relation between response strength or latency, and orientation and luminance contrast (*A, B*) Surfaces for data from one neuron. (*C, D*) Surfaces for the model of those data. The response strength is closely related to changes in orientation but shows only a small decrease as luminance contrast decreases. Latency, however, is nearly unaffected by orientation at high contrast, and the effect of orientation becomes large only at very low luminance contrast. At low contrast, there is frequently no response at nonoptimal orientations.

equations we use luminance contrast. We are now working on extensions that will provide the connection between texture differences and luminance contrast.

RELATION OF CODES WITHIN NEURONAL POPULATIONS

To this point we have discussed the structure of individual neuronal codes. However, as pointed out in the introduction, we must know how neurons act together within populations. Intuitively, it seems plausible, indeed even likely, that neurons that are located in close proximity receive many inputs in common. This appears to be the governing principle for retinotopic organization. We can surmise that the same rule is in effect for inferior temporal cortex because inferior temporal cortex neurons found near each other respond to members of the same set (Fujita et al., 1992; Gawne and Richmond, 1993; Tanaka, 1993). This grouping of neurons is the basis for the idea of columnar organization.

We can ask how much information about the stimulus is shared by neighboring neurons within the cortex. But what stimulus set do we use? We desire a stimulus set that reveals (1) all possible responses for each neuron (i.e., the full dynamic

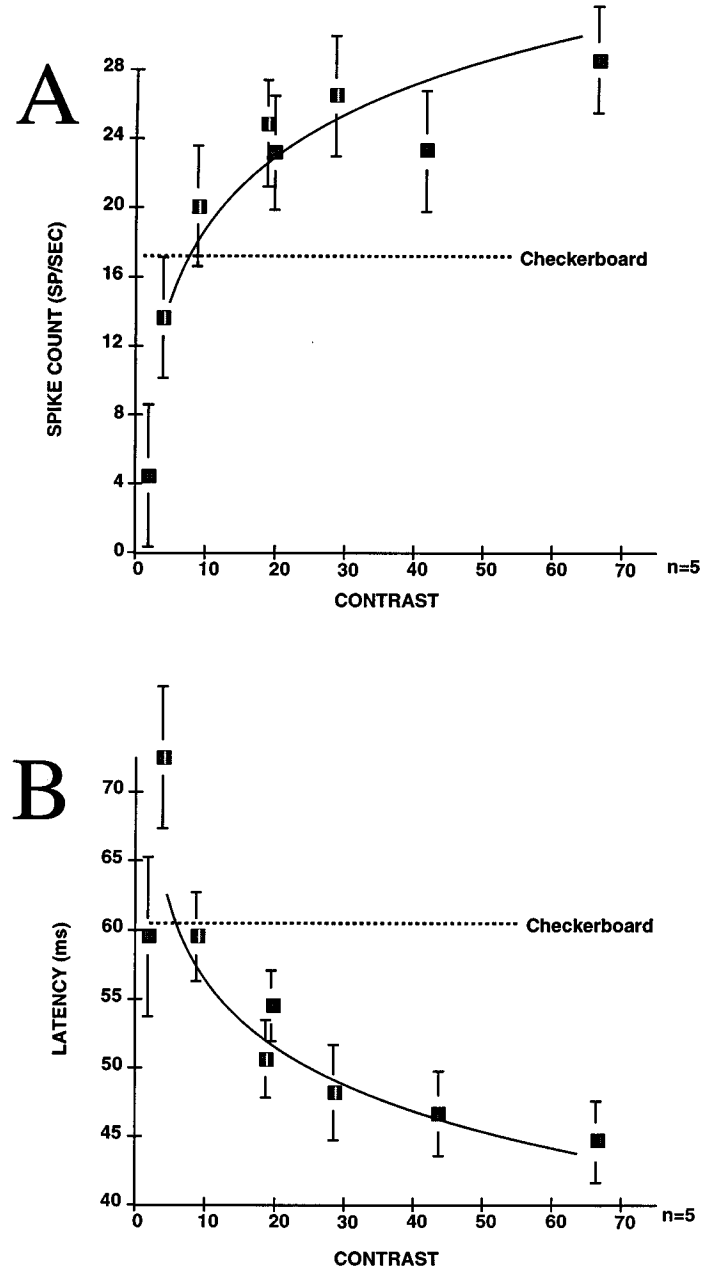


Figure 3.9 Response to the optimally oriented bar: (A) increase in response strength as a function of contrast and (B) decrease in latency as a function of contrast. The response to the optimally oriented texture defined bar is shown on both curves; it is equivalent to a bar of about 10% contrast in both cases.

range of the neuron) and (2) all possible response combinations across the pair of neurons. In inferior temporal cortex, the rules for determining neuronal selectivity are not known. Thus, it is difficult to make any a priori prediction about how the responses of one neuron relate to the responses of its neighbors.

In both V1 and inferior temporal cortex, we presented a set of two-dimensional black and white patterns based on Walsh functions, a set we have used before (Fig. 3.10A) (Richmond et al., 1987, 1990; Gawne and Richmond, 1993; Gawne et al., 1996b). In inferior temporal cortex the stimuli were centered at the point of fixation. In V1 we determined the optimal position, size, and orientation of the bar, and we centered the stimuli on the receptive field of one of the two neurons. In V1, the largest response to Walsh patterns of the index neuron was on average significantly larger ($49.5 \text{ spikes/250 ms} \pm 3.2$, SEM, $n = 26$) than the largest response to the bars (39.4 ± 5.7), even though we worked hard to find the optimal bar (Fig. 3.10B). Thus, the Walsh patterns elicited a significantly larger dynamic range than the bars we used.

The data from the pairs of neurons were analyzed in three ways:

How much of the signal variance or noise variance of one neuron could be predicted by its neighbor?

How much information about the stimulus is carried in common by neighboring neurons?

Are these related to the classical spike cross-correlogram?

The signal is taken to be the average response to each stimulus. The correlation between the average responses to the stimulus set across the neuronal pairs accounted for 19% of the variance in V1 and 22% in inferior temporal cortex. When the stimulus set was restricted to oriented bars, the variance accounted for in the regression of the signals was greater, 40%. When two neurons have correlated signals, they carry the same information, and unless we are trying to remove noise that is introduced by those neurons, having neurons carry the same information seems wasteful. Therefore, it is important to use a stimulus set that shows the full encoding capability of the neurons. Our results seem to show that this set of oriented bars yields a substantial underestimate of the potential independence of the responses of neighboring neurons. Our results raise the strong possibility that although V1 neurons are very sensitive to orientation, orientation is not the basis for encoding information about two-dimensional, black and white patterns. The amount of variance in the noise that could be accounted for in one neuron by using the responses of the other was 5.5% in both V1 and inferior temporal cortex. Thus, the correlation of the noise is considerably less than the correlation in the signal.

Using the neural network we asked whether the following equation held:

$$I(S;A,B) = I(S;A) + I(S;B) \quad (3.5)$$

Using the results from the neural network, we calculated the ratio:

$$\frac{I(S;A,B)}{I(S;A) + I(S;B)} \quad (3.6)$$

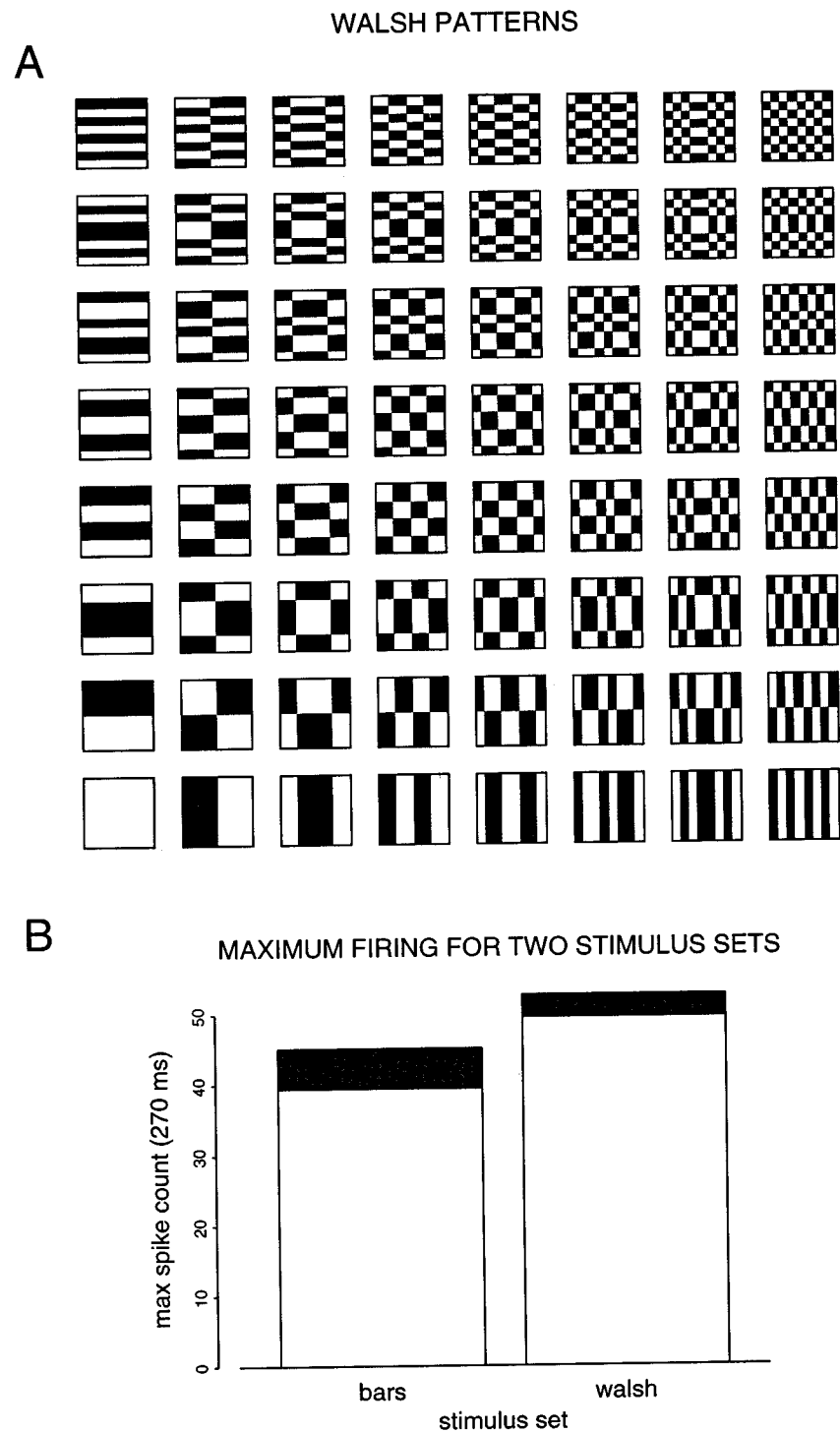
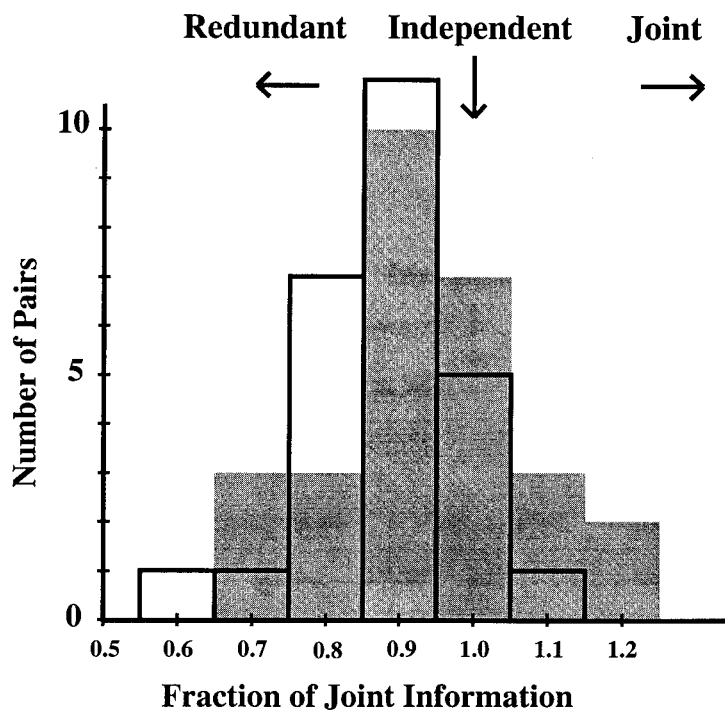


Figure 3.10 The Walsh stimulus set of 64 orthogonal stimuli mapped onto an 8×8 resolution grid. The contrast-reversed set was also used. The stimuli were centered at the center of the optimal black or white bar, whichever elicited the bigger response. (B) The average maximum firing rate for 26 neurons tested with both Walsh stimuli and bars; solid segments show the standard errors of the mean.

Figure 3.11, which sets the ratio of the information carried by the two neurons taken separately against the ratio taken together, shows that the amount of information carried in common is about 10% when the measure just described is applied.

Finally, we looked at the classical spike cross-correlograms. For the limited size samples we had, there was no significant predictive relation between the spike cross-correlograms and the response variance in number of spikes that could be predicted about one neuron's responses when the other neuron's responses were known.

The values of the noise correlation are similar to those reported from middle temporal (MT) cortex by Zohary et al. (1994). Gochin et al. (1991) recorded pairs of neurons in IT cortex, also. They concentrated more on the cross-correlogram structures, but they did report the correlations across response histograms from neuronal pairs. They used a small stimulus set and a quite different method for



Striate: n=26, two monkeys

IT: n=28, two monkeys

Figure 3.11 Information ratios (see text) for pairs of both V1 complex cells and IT neurons. If the information from the two neurons taken together had been equal to the sum of the information from each taken alone, the ratio would have been 1.0. If both neurons carried exactly the same information the ratio would have been 0.5. If there had been a joint code (e.g., a difference code), the ratio would have been larger than 1. The histograms were centered at about 0.9, with only a few pairs having ratios slightly above one.

calculating the correlations. They report the correlation coefficients r rather than the r^2 . The average value for r^2 (assuming r^2 can be gotten by squaring r), was lower than ours, on the order of 0.1 for neurons recorded from one electrode, and even lower for neurons recorded from two electrodes. Thus, it appears from all these measurements that adjacent neurons have only a very modest degree of correlation in signals and an even smaller amount of correlation in the noise.

Tanaka (1993) also pointed out that although the neurons in a local region of IT cortex all respond to stimuli in the same limited class, the pattern of responses of the separate neurons to the different members of this general class of stimuli appeared nearly independent. Tanaka's results in IT cortex are essentially the same as our results in both IT and striate cortex, where neurons located close to each other share sensitivity to a general class of stimuli (objects of related form in IT; bars or edges of the same general location in space and/or orientation in striate), but exposure to a larger and more varied set of stimuli reveals a much broader spectrum of sensitivity.

Problems Given Data from Large Neuronal Arrays

The discussions above keep alluding to problems in assessing the results of experiments involving data from many neurons. As the number of neurons goes up, the difficulty in estimating the joint response distributions accurately increases, as well. Therefore, it would be very useful to know whether simple models of the distribution can be formed. In recent work we have shown that the noise around each mean spike count is modeled well by a Gaussian (Gershon et al., 1996). In addition, it is well known that the response variance rises as the mean response strength increases (this must be true because of the floor effect—there can be no negative firing rates). Consider the information processing by a pair of neurons. Denote the total response strength by R and noise strengths by N for both neurons. Assume that the correlations between the responses of two neurons and between the noise strengths are ρ_r and ρ_n , respectively, and that the response and noise are both Gaussian. The amount of information each neuron can individually convey is

$$I = \frac{1}{2} \ln \frac{R}{N} \quad (3.7)$$

Jointly, two neurons (for convenience, assume that R/N is the same for both) convey

$$I_{\text{joint}} = \frac{1}{2} \ln \frac{R^2(1 - \rho_r^2)}{N^2(1 - \rho_n^2)} \quad (3.8)$$

If we define the degree of redundancy as

$$D_r = \frac{2I - I_{\text{joint}}}{I} \quad (3.9)$$

then

$$D_r = \frac{\ln\left(\frac{1 - \rho_n^2}{1 - \rho_r^2}\right)}{\ln\left(\frac{R}{N}\right)} \quad (3.10)$$

The total response R includes N and thus is always bigger than N . If $\rho_n^2 < \rho_r^2$, D_r is positive and there is redundancy. This is the case we intuitively expect. When $\rho_n^2 = \rho_r^2$, there is no redundancy. When $\rho_n^2 > \rho_r^2$, the two neurons carry information in a joint code. The second and third cases show that two neurons can be correlated without being redundant. Equation 3.8 shows that correlated signals decrease the total information that is transmitted and that correlated noise actually increases the total information. Taking advantage of the correlated noise requires use of a joint code. (*Hint*: Think about differential recording; the subtraction removes correlated noise.) Equation 3.10 shows that the full correlational structure (i.e., both for the response and for its noise) must be known before any conclusions regarding redundancy can be made from signal and noise measurements.

The measured signal-to-noise ratio was 1.25, the average correlation between noise of different neurons is about 0.055 in both V1 and IT cortices, and the signal correlation between two adjacent neurons is about 0.20. The information was about 20% redundant. From the measured values of the averaged transmitted information I , ρ_s , and ρ_n , D_r is predicted to be 19%, which is virtually the same as the experimental value of 20%.

NEURONAL CODES AND CORTICAL ORGANIZATION

This chapter started by listing two important ideas about cortical processing. One is that neurons are tuned, with one particular stimulus eliciting the largest response and small changes in this “optimal” stimulus causing more or less regular decreases in the response strength. However, neurons appear to have responses that are multidimensional, with the different aspects of a neuronal response specialized to carry information about different aspects of a stimulus. At a minimum, the idea of a tuning curve must be extended. In simple situations, where identifiable aspects of the response can be easily related to different aspects of the stimulus (e.g., a response latency closely related to the stimulus contrast and a response strength closely related to stimulus orientation), the extension is straightforward: a set of two-dimensional tuning surfaces replaces a single one-dimensional tuning curve (see eq. 3.10). In other situations it is not clear how one would go about representing the tuning of a neuron. For example, in the results obtained when the Walsh patterns were used as stimuli, it is not clear what aspects of these complex two-dimensional patterns are being represented in the temporal patterns of excitation and inhibition we see in the neuronal responses, so we cannot as yet draw tuning surfaces for these conditions. In either case we are faced with the task of determining how these multivariate codes are interpreted by the rest of the nervous system.

The second principle of cortical processing to be addressed is that neurons with similar stimulus or motor sensitivities tend to be grouped in columns where the processing is similar. First Mountcastle (1957) and then Hubel and Wiesel (1962) proposed that cortex is divided into functional columns, and the neurons within a

column have similar response properties. However, we now find that the responses of adjacent neurons in cortex share only about 10% of their stimulus-related information. How can these two seemingly different results be reconciled?

Consider a small group of neurons in cortex (e.g., the neurons in a column). Assume that these neurons all receive inputs that have something in common—in primary visual cortex, all the inputs to a particular column will tend to carry information about the same region of visual space; in other areas such as IT cortex, the common factor to the inputs is not known, but presumably there is some pattern. All by itself the limited nature of the inputs to a given region will tend to make the neurons in this region sensitive to particular classes of stimuli (e.g., in striate cortex, stimuli in a particular part of visual space). However, this constraint says nothing about the relationships between the neurons and the large numbers of stimuli that belong to the same general class. By having different patterns of connections from the same set of inputs, these neurons can have response properties that are independent of each other within the same broad class of stimuli. The simplest possible case is illustrated by a cartoon (Fig. 3.12). Each neuron computes a different function of the inputs. We would require that each neuron compute the function

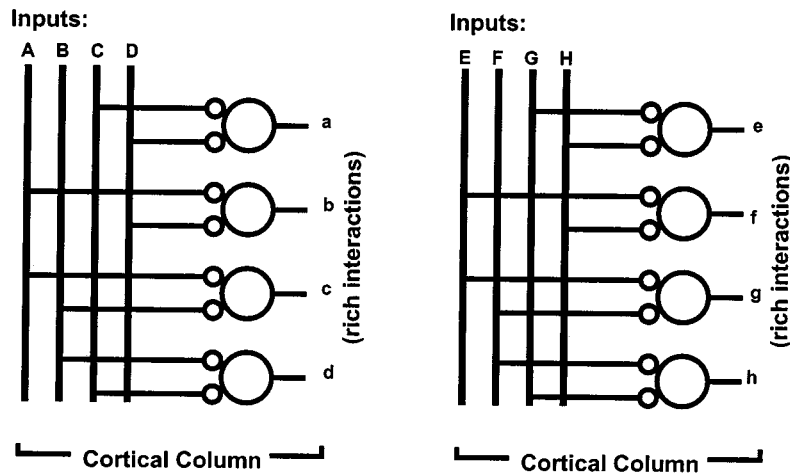


Figure 3.12 Highly simplified cartoon diagram of what might be happening in cortical columns. Different columns get different sets of input fibers (ABCD for the column on the left and EFGH for the column on the right). However, neurons within each individual column are connected to these afferents in different ways, here schematized by different patterns of excitatory connections. Inhibitory connections and/or processing by interneurons could also have the same effect. If different functions of the afferents were calculated, the neurons within each column would have response properties nearly independent of each other across the entire set of input patterns that could be supported by its set of inputs. The independent patterns of activity by the neurons within a column are the basis for interactions that maximize the use of the relatively great density of local connections. Note that the use of very simplified stimulus sets would fail to uncover this organization. If one probed this structure with a stimulus set having only two conditions, one where all of ABCD were “on” and all of EFGH were “off,” and the other condition where all of ABCD were off and all of EFGH were on, one could determine only that the neurons in any single column all tend to respond to the same stimuli; the diversity of behavior within a column could not be observed.

of the inputs using strictly local connections and that the wiring changes from neuron to neuron be as few as possible, consistent with achieving relatively independent processing of local inputs.

Why should neurons in a column have such diverse response properties? Why not simply have all the neurons code for a specific stimulus feature redundantly, thus giving a single very precise and robust message? We don't know the answer to this question just yet, but we can speculate. Consider the problem of correlated signals. To the extent that the noise on a neuron's output is due to the intrinsic unreliability of the neuron itself, then averaging the outputs of multiple identical neurons would be useful. It would be expensive in that the amount of signal processing possible would be very limited (noise decreases by \sqrt{n} , where n is the number of elements). Furthermore, recent results show that neurons may not be very noisy, making averaging an even weaker candidate as a signal processing strategy (Mainen and Sejnowski, 1995).

Another way to view the problem is to consider the great cost of connections within the nervous system. As with the design of large computer systems, the number and length of connections may be a greater constraint on the operation of a complex nervous system than the number of active processing elements (Van Essen, 1997). It is obvious from the geometry that most connections will be local and long-range connections will be in some sense expensive to the system and fewer in number. That this constraint would cause regions of cortex that are processing information about related stimuli to lie in close proximity to each other has been pointed out before. What is more subtle is the notion that while one might well want neurons that are carrying information about stimuli that are in some sense related to lie next to each other, it would be incredibly wasteful for neurons carrying strongly correlated information to lie next to each other. The reason should be obvious: if the responses of two neurons are correlated, there is no need to connect them. To maximize the utilization of local connections, there should be strong pressure to maximize the diversity of response properties in a local region, to maximize, in turn, the ability of local groups of neurons to interact with each other over their comparatively rich set of local connections. If for some reason the nervous system did have many neurons carrying duplicated signals, it might make more sense for these identical neurons to be as widely distributed as possible, to maximize their ability to interact via local connections with other neurons that are carrying different information.

REFERENCES

- Abeles, M., Bergman, H., Margalit, E., and Vaadia, E. (1993) Spatiotemporal firing patterns in the frontal cortex of behaving monkeys. *J. Neurophysiol.* 70:1629–1638.
- Ahmed, N., and Rao, K. R. (1975) *Orthogonal Transforms for Digital Signal Processing*. Berlin: Springer-Verlag.
- Barlow, H. B. (1972) Single units and sensation: A neuron doctrine for perceptual psychology? *Perception* 1:371–394.
- Bialek, W., Rieke, F., de Ruyter van Steveninck, R. R., and Warlund, D. (1991) Reading a neural code. *Science* 252:1854–1857.
- Cover, T. M., and Thomas, J. A. (1991) *Elements of Information Theory*. New York: Wiley.

- de Ruyter van Steveninck, R., and Bialek, W. (1988) Real-time performance of a movement-sensitive neuron in the blowfly visual system: Coding and information transfer in short spike sequences. *Proc. R. Soc. London B* 234:379–414.
- Eckhorn, R., Bauer, R., Jordan, W., Brosch M., Kruse W., Munk, M., and Reitboeck, H. J. (1988) Coherent oscillations: A mechanism of feature linking in the visual cortex? Multiple electrode and correlation analyses in the cat. *Biol. Cybern.* 60:121–131.
- Fujita, I., Tanaka, K., Ito, M., and Cheng, K. (1992) Columns for visual features of objects in monkey inferotemporal cortex. *Nature* 360:343–346.
- Gawne, T. J., and Richmond, B. J. (1993) How independent are the messages carried by adjacent neurons in inferior temporal cortex? *J. Neurosci.* 7:2758–2771.
- Gawne, T. J., Richmond, B. J., and Optican, L. M. (1991) Interactive effects among several stimulus parameters on the responses of striate cortical complex cells. *J. Neurophysiol.* 62:379–389.
- Gawne, T. J., Kjaer, T. W., and Richmond, B. J. (1996a) Latency: Another potential code for feature binding in striate cortex. *J. Neurophysiol.* 76:1356–1360.
- Gawne, T. J., Kjaer, T. W., Hertz, J. A., and Richmond, B. J. (1996b) Adjacent visual cortical complex cells share about 20% of their stimulus-related information. *Cereb. Cortex* 6:482–489.
- Georgopoulos, A. P., Kettner, R. E., and Schwartz, A. B. (1988) Primate motor cortex and free arm movement to visual targets in three-dimensional space. II. Coding of the direction of movement by a neuronal population. *J. Neurosci.* 8:2928–2937.
- Gershon, E. D., Latham, P. E., Jin, G.-X., and Richmond, B. J. (1996) Stimulus-elicited neuronal responses in striate and inferior temporal cortices are well-described by a Gaussian distribution. *Soc. Neurosci. Abstr.* 22:1612.
- Gochin, P. M., Miller, E. K., Gross, C. G., and Gerstein, G. L. (1991) Functional interactions among neurons in inferior temporal cortex of the awake macaque. *Exp. Brain Res.* 84:505–516.
- Golomb, D., Hertz, J., Panzeri, S., Treves, A., and Richmond, B. (1997) How well can we estimate the information carried in neuronal responses from limited samples? *Neural Computation*, 9:649–665.
- Gray, C. M., and Singer, W. (1989) Stimulus-specific neuronal oscillations in orientation columns of cat striate cortex, *Proc. Natl. Acad. Sci. U.S.A.* 86:1698–1702.
- Heller, J., Hertz, J. A., Kjaer, T. W., and Richmond, B. J. (1995). Information flow and temporal coding in primate pattern vision. *J. Comput. Neurosci.* 2:175–193.
- Hubel, D. H., and Wiesel, T. N. (1962) Receptive fields, binocular interaction and functional architecture in cat's visual cortex. *J. Physiol.* 160:106–154.
- Kjaer, T. W., Hertz, J. A., and Richmond, B. J. (1994) Decoding cortical neuronal signals: Network models, information estimation and spatial tuning. *J. Comput. Neurosci.* 1:109–139.
- Lestienne, R., and Strehler, B. L. (1987) Time structure and stimulus dependence of precisely replicating patterns present in monkey cortical neuronal spike trains. *Brain Res.* 437:214–238.
- Mainen, Z. F., and Sejnowski, T. J. (1995) Reliability of spike timing in neocortical neurons. *Science* 268:1503–1506.
- Meister, M. (1996) Multineuronal codes in retinal signaling. *Proc. Natl. Acad. Sci. U.S.A.* 93:609–614.
- Miller, J. P., and Theunissen, F. (1995) Temporal encoding in nervous systems: A rigorous definition. *J. Comput. Neurosci.* 2:149–162.
- Movshon, J. A., Thompson, I. D., and Tolhurst, D. J. (1977a) Spatial summation in the receptive field of simple cells in the cat's striate cortex. *J. Physiol.* 283:53–77.
- Movshon, J. A., Thompson, I. D., and Tolhurst, D. J. (1977b) Receptive field organization of complex cells in the cat's striate cortex. *J. Physiol.* 283:79–99.

- Mountcastle, V. B. (1957) Modality and topographic properties of single neurons of cat's somatic sensory cortex. *J. Neurophysiol.* 20:408–434.
- Optican, L. M., and Richmond, B. J. (1987) Temporal encoding of two-dimensional patterns by single units in primate inferior temporal cortex. III. Information theoretic analysis. *J. Neurophysiol.* 57:162–178.
- Optican, L. M., Gawne, T. J., Richmond, B. J., and Joseph, P. J. (1991) Unbiased measures of information transmission and channel capacity in neurons. *Biol. Cybern.* 65:305–310.
- Panzeri, S., and Treves, A. (1996) Analytical estimates of limited sampling bias in different information measures. *Network* 7:87–107.
- Richmond, B. J., and Optican, L. M. (1987) Temporal encoding of two-dimensional patterns by single units in primate inferior temporal cortex. II. Quantification of response waveform. *J. Neurophysiol.* 57:147–161.
- Richmond, B. J., and Optican, L. M. (1990) Temporal encoding of two-dimensional patterns by single units in primate visual cortex. II. Information transmission. *J. Neurophysiol.* 64:370–380.
- Richmond, B. J., Optican, L. M., Podell, M., and Spitzer, H. (1987) Temporal encoding of two-dimensional patterns by single units in primate inferior temporal cortex. I. Response characteristics. *J. Neurophysiol.* 57:132–146.
- Richmond, B. J., Optican, L. M., and Spitzer, H. (1990) Temporal encoding of two-dimensional patterns by single units in primate visual cortex. I. Stimulus–response relations. *J. Neurophysiol.* 64:351–369.
- Shatz, C. J. (1996) Emergence of order in visual development. *Proc. Natl. Acad. Sci. U.S.A.* 93:602–608.
- Spitzer, H., and Hochstein, S. (1985) Simple- and complex-cell response dependences on stimulation parameters. *J. Neurophysiol.* 53:1244–1265.
- Tanaka, K. (1993) Neuronal mechanisms of object recognition. *Science* 262:685–688.
- Tovée, M. J., Rolls, E. T., Treves, A., and Baylis, R. P. (1993) Information encoding and the responses of single neurons in the primate temporal visual cortex. *J. Neurophysiol.* 70:640–654.
- Vaadia, E., Haalman, I., Abeles, M., Bergman, H., Prut, Y., Slovin, H., and Aertsen, A. (1995) Dynamics of neuronal interactions in monkey cortex in relation to behavioural events. *Nature* 373:515–518.
- Van Essen, D. C. (1997) A tension-based theory of morphogenesis and compact wiring in the central nervous system. *Nature* 385:313–318.
- Zohary, E., Shadlen, M. N., and Newsome, W. T. (1994) Correlated neuronal discharge rate and its implications for psychophysical performance. *Nature* 370:140–142.